SEMINARIO DI ANALISI MATEMATICA
DIPARTIMENTO DI MATEMATICA DELL'UNIVERSITA' DI BOLOGNA

S. GRAFFI

TEORIA CANONICA DELLE PERTURBAZIONI IN MECCANICA QUANTISTICA
I parte

2 APRILE 1987

1. TEORIA CANONICA DELLE PERTURBAZIONI

Sia M una superficie analitica senza bordo di dimensione ℓ, e sia T*M il suo fibrato cotangente. Indichiamo con (p,q) le coordinate (locali) su T*M e sia $(p,q) \to H(p,q) \in C^\infty(T^*M\,;\,R)$. Nel seguito, considereremo il flusso Hamiltoniano associato alle equazioni canoniche:

$$\dot{p} = -\nabla_q H \quad , \quad \dot{q} = \nabla_p H \tag{1.1}$$

1.1. Varie nozioni di integrabilità

Il flusso Hamiltoniano (1.1) o, equivalentemente, l'Hamiltoniana H(p,q) si chiama *integrabile* sull'aperto $W \subset T^*M$ (più precisamente, *canonicamente integrabile*) se e solo se esiste una trasformazione completamente canonica analitica da W a $V \times T^\ell$ (V aperto in R^ℓ ; T^ℓ toro ℓ-dimensionale) $(A,\phi) = C(p,q)$, $A \in R^\ell$, $\phi \in T^\ell$, tale che

$$H(C^{-1}(p;q)) = h(A), \quad (A,\phi) \in V \times T^\ell \tag{1.2}$$

Le variabili canoniche (A,ϕ) vengono dette variabili azione-angolo.

A proposito di questa definizione, osserviamo

(a) Nelle coordinate (A,ϕ) le equazioni del moto sono:

$$\dot{A} = 0 \quad , \quad \dot{\phi} = \omega(A) \quad , \quad \omega(A) = \nabla_A h(A) \tag{1.3}$$

Pertanto le componenti di A sono ℓ integrali primi, e il moto appare come una rotazione uniforme su T^ℓ : $\phi \to w(A)t + \phi$. In altre parole tutti i moti in W sono quasi periodici.

(b) Ricordiamo che una trasformazione $(p_1,q_1) = C(p,q)$ da W a W_1 si dice completamente canonica se conserva la 2-forma $\sum_{i=1}^{\ell} dp_i \wedge dq_i$: $\sum_{i=1}^{\ell} dp_i \wedge dq_i =$

$$= \sum_{i=1}^{\ell} dp_i^1 \; dp_i^1 .$$

c) La trasformazione canonica $(A,\phi) = C(p,q)$ che integra H, cioè che la coniuga canonicamente all'Hamiltoniana $h(A)$, ha l'effetto di definire una *foliazione* canonica di W in tori invarianti per il moto. Tale foliazione dello spazio delle fasi è descritta dalle equazioni della trasformazione canonica stessa in fatti

$$(1.5) \qquad (p,q) = C^{-1}(A,\phi) \quad , \quad \phi \in T^{\ell}$$

definisce un toro (parametrizzato da A) $\mathscr{I}_A$ invariante per il moto per quanto detto sopra; al variare di A in V si descrive W.

(d) Osserviamo che, data l'invarianza delle parentesi di Poisson per trasformazioni canoniche, le azioni A_i non solo sono integrali primi, ma sono anche in involuzione:

$$(1.6) \qquad \{A_i,A_j\} \equiv \sum_{k=1}^{\ell} \left(\frac{\partial A_j}{\partial q_k}\frac{\partial A_i}{\partial p_k} - \frac{\partial A_i}{\partial q_k}\frac{\partial A_j}{\partial p_k}\right) = 0.$$

Pertanto se un sistema è canonicamente integrabile esso ammette n integrali primi in involuzione e l'energia (cioè la funzione hamiltoniana stessa) può essere espressa in funzione di questi. Di solito sono queste proprietà quelle che si usano per definire un sistema integrabile: la definizione più comune di sistema integrabile è infatti la seguente: il sist. Hamiltoniano di Hamiltoniana analitica H su T^*M è integrabile su $W \subset T^*M$ se esso ammette ivi ℓ integrali primi analitici I_h, $h = 1,\dots,\ell$, indipendenti e in involuzione, tali che l'energia è una funzione analitica degli integrali primi: $h = h(I_1,\dots,I_\ell)$.Che la seconda definizione sia equivalente alla prima è conseguenza di un teorema celebre.

Teorema (Liouville-Arnold). Supponiamo che il sistema Hamiltoniano H su T^*M ammetta sull'aperto $W \subset T^*M$ ℓ integrali primi analitici $I = (I_1,\dots,I_\ell)$ in involuzione $\{I_i, I_j\} = 0$ e che l'energia H sia funzione di questi. Supponiamo inoltre che $\exists I_0$ tale che le superficie I_i = cost, $i = 1,\dots,\ell$, siano compatte e senza bordo per I sufficientemente vicino a I_0. Allora il sistema (H,W) è canonicamente integrabile in un opportuno intorno di $I = I_0$.

Per la dim., si vedano le note bibliografiche. I punti essenziali sono il dimostrare (1) che le superfici I_i = cost. , $i = 1,\dots,\ell$, sono tori a ℓ dimensioni (2) che tali superfici sono varietà Lagrangiane, cioè tali che la forma differenziale $\sum p_i dq_i$ è ivi localmente integrabile. Ciò permette di definire le variabili di azione.

1.2. La teoria formale delle perturbazioni

Gli esempi di sistemi integrabili nel senso predetto non sono molti: problema di Keplero, moto geodetico sull'ellissoide, giroscopio, e pochi altri, fra cui gli oscillatori armonici: qui $M = R^\ell$, $T^*M = R^{2\ell}$,

$$H(p,q) = \frac{1}{2}\sum_{i=1}^{\ell}(p_i^2 + \omega_i^2\, q_i^2) \tag{1.7}$$

$$C^{-1}(A,\ell) := \begin{cases} p_i = -\sqrt{2A_i\omega_i}\ \sin\phi_i \\ q_i = \sqrt{\dfrac{2A_i}{\omega_i}}\ \cos\phi_i \end{cases} \qquad i = 1,\dots,\ell \tag{1.8}$$

$$C(A,\ell) := \begin{cases} A_i = (p_i^2 + \omega_i^2 q_i^2)/2\omega_i \\ \phi_i = -\operatorname{arctg}\left(\dfrac{\omega_i p_i}{q_i}\right) \end{cases} \tag{1.9}$$

Si verifica immediatamente che C^{-1} mappa $V \times T^\ell$, $V = R_+$, su $(R^2\backslash\{0\})^\ell$, e

$$H(C^{-1}(A,\phi)) = \sum_{i=1}^{\ell} \omega_i A_i \equiv \langle\omega,A\rangle.$$

Sono molti i sistemi hamiltoniani che, però, possono essere considerati quasi integrabili, cioè che possono essere scritti sotto la forma $H_o(p,q) + \varepsilon V(p,q)$ dove H_o è integrabile nel senso precedente. L'esempio più classico è costituito dal modello standard del sistema planetario, il problema degli n corpi in cui si considerino le attrazioni interplanetarie come perturbazioni delle attrazioni fra il sole ed ogni singolo pianeta.

Poiché la trasformazione $C(p,q) = (A,\phi)$ che integra H_o è completamente canonica, possiamo considerare $H(p,q)$ nelle coordinate (A,ϕ). L'Hamiltoniana potrà cioè scriversi così:

$$(1.10) \qquad H(A,\phi) = h_o(A) + \varepsilon V(A,\phi) \quad ; \quad h_o(A) \equiv H_o(C^{-1}(A,\phi))$$

$$V(A,\phi) \equiv V(C^{-1}(A,\phi))$$

$H(A,\phi)$ essendo definita su $V \times T^\ell$ ed ivi analitica. Lo scopo della teoria canonica delle perturbazioni è la ricerca di una trasformazione canonica $C_\varepsilon(A,\phi) = (A_1,\phi_1)$ da $V \times T^\ell$ in $V_1 \times T^\ell$, $V_1 \subset R^\ell$, analitica in (A,ϕ,ε), tale da rimuovere la dipendenza da ϕ sull'Hamiltoniana trasformata ad ogni ordine in ε, cioè tale che:

$$(1.11) \qquad H(C_\varepsilon^{-1}(A_1,\phi_1)) = h_1(A_1,\varepsilon) + \varepsilon^{n+1} f_{n+1}(A_1,\phi_1)$$

n positivo arbitrario.

La forma (1.11) permette immediatamente di stabilire che i moti di (1.10) si manterranno periodici per tempi dell'ordine di ε^{-n} per $\varepsilon \to 0$.

L'algoritmo formale è il seguente: si cerca una funzione generatrice $S_\varepsilon(A_1,\phi)$, analitica in $V \times T^\ell \times \{\varepsilon: |\varepsilon|<c\}$ per la trasformazione canonica, cioè tale che

$$(1.12) \qquad \begin{cases} A = \nabla_\phi S_\varepsilon(A_1,\phi) \\ \phi_1 = \nabla_{A_1} S_\varepsilon(A_1,\phi) \end{cases}$$

Modulo le condizioni di invertibilità (discusse in seguito) le (1.12) definiscono, come è noto, C_ε.

Per determinare S_ε si usa l'equazione di Hamilton-Jacobi, ottenuta richiedendo che l'immagine di $H(A,\phi)$ nelle nuove variabili sia una funzione indipendente dagli angoli fino all'ordine ε^{n+1}.

Inserendo la prima di (1.12) in (1.10) e imponendo l'indipendenza dagli angoli fino all'ordine ε^{n+1} si ottiene l'equazione di Hamilton-Jacobi perturbativa:

$$(1.13) \qquad h_0(\nabla_\phi S_\varepsilon(A_1,\phi)) + \varepsilon V(\nabla_\phi S_\varepsilon(\cdot),\phi) = h_\varepsilon(A_1,\varepsilon) + \varepsilon^{n+1} f_{n+1}(A_1,\phi)$$

che deve determinare sia S_ε che (h_1,f_{n+1}).

Si cerca ora una soluzione di (1.13) sotto forma di serie di potenze, a priori formale, cercando di determinarne ricorsivamente i coefficienti:

$$(1.14) \qquad S_\varepsilon(A_1,\phi) = S_0(\cdot) + \varepsilon\, S_1(\cdot)+\ldots,\ h_\varepsilon(A_1) = h_0(A_1)+\varepsilon\, h^1,\ldots,$$

Poiché h_0 è integrabile, S_0 dovrà essere la funzione generatrice della trasformazione canonica identica:

$$(1.15) \qquad S_0(A_1,\phi) = \langle A_1,\phi\rangle.$$

Inserendo (1.14) e (1.15) in (1.13), e limitandoci per il momento a sviluppare ambo i membri fino al 1° ordine incluso si avrà:

$$(1.16) \qquad h_0(A_1)+\varepsilon\langle\omega(A_1),\nabla_\phi S_1(A_1,\phi)\rangle+\varepsilon\, V(A_1,\phi) = h_0(A_1) + \varepsilon h^1(A_1)$$

da cui l'equazione di Hamilton-Jacobi al 1° ordine

$$(1.17) \qquad \langle\omega(A_1),\nabla_\phi S_1(A_1,\phi)\rangle + V(A_1,\phi) = h^1(A_1)$$

che si risolve introducendo lo sviluppo di Fourier

$$(1.18)\qquad S_1(A_1,\phi) = \sum_{\nu\in Z^\ell} S_1^\nu(A_1)\, e^{i<\nu,\phi>} \;;\; V(A_1,\phi) = \sum_{\nu\in Z^\ell} V_\nu(A_1) e^{i<\nu,\phi>}$$

Pertanto la (1.17) scritta sui coefficienti di Fourier diventa

$$(1.19)\qquad i<\omega(A_1),\nu>\, S_1^\nu(A_1) + V\,(A_1) = h^1(A_1)$$

da cui $h^1(A_1) = V_o(A_1)$, $S_1^\nu(A_1) = -i\,\dfrac{V_\nu(A_1)}{<\omega(A_1),\nu>}$, $\nu \neq 0$ $\quad S_1^o(A_1) = 0$, e la serie $-i\Sigma\,\dfrac{V_\nu(A_1)}{<\omega(A_1),\nu)}\, e^{i<\nu,\phi>}$ convergerà per tutti gli A_1 tali che $<\omega(A_1),\nu)$ non tende a 0 troppo rapidamente per $|\nu| \to \infty$. Posponendo al paragrafo successivo una breve disamina del fenomeno dei "piccoli denominatori", essenziale per discutere la convergenza della serie di Fourier, notiamo che l'esistenza stessa della serie richiede che sia soddisfatta la *condizione di non risonanza*

$$(1.20)\qquad <\omega(A_1),\nu> = 0 \Leftrightarrow \nu = 0$$

I punti $A\in V$ tali che $\exists\bar\nu \neq 0$ con $<\omega(A),\bar\nu> = 0$ sono detti risonanze di h_o di ordine ν.

Escludendo dunque da V tutti i punti risonanti e dei loro opportuni intorni se un $<\omega(A),\nu>$ diventa "troppo piccolo" l'algoritmo di eliminazione degli angoli può essere applicato a tutti gli ordini in ε. Infatti inserendo ancora (1.14) in (1.13) e imponendo per $1\leq k\leq n$, che il 1° membro ha una funzione indipendente da ϕ si ottengono le equazioni ricorsive

$$(1.21)\qquad <\omega(A_1),\nabla_\phi S_k(A_1,1)> + N_k(A_1,\phi) = h_k(A_1)$$

dove N_k è un polinomio in $\nabla_\phi S_j\,(A_1,\phi)$, $j=1,\dots,h-1$, con coefficienti proporzionali a $\frac{\partial^{|a|}}{\partial A_1^a} V(A_1,\phi)$, $|a| < k$ o a $\frac{\partial^{|a|}}{\partial A_1^a} h_0(A)$, $|a| \leq k$.

La (1.21) implica chiaramente $h_k(A_1) = \overline{N_k(A_1,\phi)} = \frac{1}{(2\pi)^\ell} \int_{T^\ell} N_k(A_1,\phi)d\phi_1\dots d\phi_\ell$ e, come sopra:

$$S_k(A_1,\phi) = \sum_{\substack{\nu\in Z^\ell \\ \nu\neq 0}} S_k^\nu(A_1)e^{i\langle\nu,\phi\rangle} \ , \ S_k^\nu(A_1) = -i\,\frac{N_k^\nu(A_1)}{\langle\omega(A_1),\nu\rangle} \tag{1.22}$$

1.3. Il teorema della forma normale di Birkhoff

Il teorema di Birkhoff riguarda il caso speciale in cui l'Hamiltoniana integrabile h_0 è un sistema di n oscillatori armonici non risonanti, $h_0(A) = \langle\omega,A\rangle$. In sostanza si rende rigoroso il metodo formale di cui sopra, mostrando che sotto opportune ipotesi i piccoli denominatori possono essere controllati e che la S_ε è veramente la funzione generatrice di una trasformazione canonica che coniuga la data Hamiltoniana con una trasformata in cui la dipendenza degli angoli può essere eliminata a tutti gli ordini. Quest'ultima Hamiltoniana prende il nome di forma normale di Birkhoff.

Consideriamo ancora l'Hamiltoniana $H_\varepsilon(A,\phi) = h_0(A) + \varepsilon V(A,\phi)$ che definendo $z = (z_1,\dots,z_\ell) = (e^{i\phi_1},\dots,e^{i\phi_\ell})$ considereremo definita su $W(\rho,\xi,V)\subset C^{2\ell}$,

$$W(\rho,\xi;V) = \{(A,z)\,|\,(A,z)\in C^{2\ell} \ , \ \exists A_0\in V \text{ tale che } |A_i-A_{0i}| < \ell\,|e^{-\xi} < |z_i| < e^{\xi} \ , \ i = 1,\dots,\ell\} \tag{1.23}$$

per $\rho>0$ e $\xi>0$ dati. Misureremo per la "grandezza" dell'Hamiltoniana H_0 e dalla sua perturbazione con le seguenti norme di funzioni olomorfe:

(1.24) $E_o = \sup_{w(\rho,\xi,v)} |\nabla_A h_o(A)|$; $\|V\|^1_{\rho,\xi} = \sup_{w(\rho,\xi;v)} (|\nabla_A V| + \frac{1}{\xi}|\nabla_\phi V|)$

L'affermazione è allora la seguente.

Teorema (Birkhoff). Sia $h_o(A) = \langle\omega,A\rangle$, dove le frequenze ω soddisfano la seguente condizione diofantina di non risonanza:

(1.25) $\exists C>0,\ \alpha>0$ tali che $|\langle\omega,\nu\rangle^{-1}| \leq C\,|\nu|^\alpha$, $\nu \neq 0$, $\nu \in Z^\ell$

$|\nu| = |\nu_1| + \dots |\nu_\ell|$.

Allora dati $n \in N$ arbitrario e $0<\delta<\xi$, le formule (1.14), (1.15) definiscono la funzione generatrice S_ε di una trasformazione analitica C_ε tale che (1.13) è soddisfatta; più precisamente:

$\exists$ una famiglia di trasformazioni completamente canoniche C_ε, C'_ε tali che:

(1) $D(C_\varepsilon) \supset V(\rho e^{-\delta}\ ,\ \xi-2\delta;V)$

(2) C_ε è analitica su $W(\cdot)$e in ε per ε sufficientemente piccolo

(3) $C'_\varepsilon\, W(\rho\ e^{-2\delta},\xi-2\delta;V) \subset W(\rho\ e^{-\delta},\xi-\delta;V)$

(4) $C_\varepsilon \circ C'_\varepsilon = Id = C'_\varepsilon \circ C_\varepsilon$ su $W(\rho\ e^{-2\delta},\xi-2\delta;V)$

(5) Sia $(A,z) = C_\varepsilon(A_1,z_1)$, $(A_1,z_1) \in W(\rho\ e^{-2\delta},\xi-2\delta;V)$.

Allora C_ε è generata da $S_\varepsilon(A_1,\phi) = \sum_{k=0}^{n} S_k(A_1,\phi)\varepsilon^k$, cioè:

(1.26) $$A = A_1 + \nabla_\phi(\varepsilon S_1 + \dots \varepsilon^n S_n)(A_1,\phi)$$

$$\phi_1 = \phi + \nabla_{A_1}(\varepsilon S_1 + \dots \varepsilon^n S_n)(A_{1,\phi})$$

(6) $$H_\varepsilon(C_\varepsilon(A_1,\phi_1)) = h_0(A_1)+\varepsilon h_1(A_1)+\ldots+\varepsilon^{n+1}f_{n+1}(A_1,\phi,\varepsilon)$$

$f_{n+1}(A_1,\phi_1,\varepsilon)$ analitica in $W(\rho e^{-2\delta},\xi-2\sigma,V)$ in (A_1,ϕ_1) e in ε sufficientemente piccolo.

Dim. (Cenno). Consideriamo la (1.21), che omettendo l'indice k potremo riscrivere come

(1.27) $$\langle\omega(A_1),\ \nabla_\phi S(A_1,\phi)\rangle = -i[N(A_1,\phi)-\bar N(A_1)]$$

dove $\omega(A_1) = \omega$ è indipendente da A_1. Per $S = S_1$, $N = N_1 = V$, l'ipotesi diofantina implica che $|S_1^\nu(A_1)| \leq G\,|\nu|^\alpha\ |V_\nu(A_1)| \leq C|\nu|^\alpha\ e^{-\xi|\nu|}$ per l'analiticité di V in $W(\rho,\xi;V)$. Pertanto anche $S_1(A_1,z)$ è analitica in tale regione. Per ricorrenza abbiamo subito che $|S^\nu(A_1)| \leq C|\nu|^{p\alpha}\ e^{-\xi|\nu|}$ per un qualche intero p, e quindi $S_k(A,z)$ esiste per ogni $k \leq u$ ed è analitica in $W(\varepsilon,\xi,V)$. Dal teorema di Cauchy seguono immediatamente le stime seguenti valide per un costante B che dipende solo da ℓ:

(1.28) $$\|S\|_{\rho e^{-\delta},\xi-\delta} \equiv \max_{W(\rho e^{-\delta},\xi,\delta,V)} |S| \leq B_\ell C\ \delta^{-\ell-\alpha}\ \|N\|_{\rho,\xi}$$

(1.29) $$\|S\|^1_{\rho e^{-\delta},\xi,\delta} \leq \|N\|_{\ell,\xi}\ B_\ell C\quad \delta^{-\ell-\alpha-\ell}\ e^{-\ell}$$

(1.30) $$\left\|\frac{\partial^2 S}{\partial A\partial\phi}\right\|_{\rho e^{-\delta},\xi-\delta} \leq BC\ \delta^{-\ell-\alpha-2}\ \|N\|_{\rho,\xi}$$

dove con N si intende sempre $N(A_1,\phi)$ privata della sua media $N = N° = N(A_1,\ell)- \bar N(A_1,\phi)$.

Partendo al solito da $N° = V(A_1,\phi) - \bar V(A_1,\ell)$ e ragionando ricorsivamente non è difficile vedere che, dato $\eta < 1$, si può prendere sempre $\varepsilon \in C$

con $|\xi|$ così piccolo tale da avere per un qualche $\delta < \delta_1 < \xi$.

$$(1.31)\qquad \|\nabla_\phi(\varepsilon S_1 + \dots \varepsilon^n S_n)(A_1,z)\|_{\rho e^{-\delta_1},\xi-\delta_1} < \eta$$

$$(1.32)\qquad \|(\varepsilon S_1+\dots\varepsilon^n S_n)(A_{1,z})\|_{\rho e^{-\delta_1},\xi-\delta_1} < \eta$$

$$(1.33)\qquad \|\frac{\partial^2}{\partial A\partial\psi}(\varepsilon S_1+\dots\varepsilon^n S_n)(A_1,z)\|_{\rho,\ell^{-\delta_1},\xi-\delta_1} < \eta$$

Se ora consideriamo le (1.26), in cui la 2ª viene riscritta in termini della variabile complessa z:

$$(1.34)\qquad z_1 = z\, e^{i\nabla_{A_1}(\varepsilon S_1+\dots\varepsilon^n S_n)(A_1,z)}$$

dalle stime precedenti segue, per il teorema olomorfo sulle funzioni implicite, che (1.26) a (1.34) possono essere esplicitate su $W(\rho e^{-2\delta},\xi-2\delta;V)$ per le funzioni ivi analitiche $\Xi(A_1,z_1)$, $\Xi_1(A,z)$, $\Delta(A_1,z_1)$, $\Delta_1(A\ ,z)$, analitiche a $\varepsilon=0$ e tali che:

$$(1.35)\qquad C_\varepsilon:\quad A = A_1 + \Xi(A_1,z_1),\ z = z_1\, e^{i\Delta(A_1,z_1)}$$

$$(1.36)\qquad C_\varepsilon^{-1}:\quad A_1 = A + \Xi_1(A\ ,z),\ z_1 = z\, e^{-\Delta_e(A\ ,z)}$$

con $C_\varepsilon^{-1}\, W(\rho e^{-2\delta},\xi-2\delta,V)\subset W(\rho e^{-\delta},\xi-\delta,V)$ così che la (6) vale su $W(\ e^{-2\delta},\xi-2\delta,V)$. Ciò conclude il cenno della prova.

Il teorema precedente mostra che la serie di Birkhoff:

$$(1.37)\qquad h(A_1,\varepsilon) = \sum_{n=i}^{\infty} h_n(A_1)\varepsilon^n$$

(1.38) $$S(A_1,\phi,\varepsilon) = \sum_{n=1}^{\infty} S_n(A_1,\phi)\varepsilon^n$$

hanno senso formalmente. Esse poi genericamente divergono (per casi particolari in cui convergono, si veda la nota bibliografica) come conseguenza del seguente celebre teorema di Poincaré:

Teorema della non-esistenza degli integrali primi uniformi (di Poincaré). (Notiamo che integrale primo uniforme = integrale primo analitico in ε nella locazione di Poincaré). Alle ipotesi precedenti di analiticità su $H(A,\phi) = h(A) + \varepsilon V(A,\phi)$ aggiungiamo la non isocronia cioè l'invertibilità della matrice $\frac{\partial^2 h}{A_i A_j}$, $A \in V$.

Allora, genericamente in V, non esistono integrali primi analitici in (A,ϕ) ed ε se non l'energia stessa.

Dim. Sia $B(A,\phi,\varepsilon)$ un integrale primo analitico su $V \times T^{\ell} \times \{\varepsilon: |\varepsilon|<\varepsilon_0\}$:

(1.37) $$B(A,\phi,\varepsilon) = B_0(A,\phi) + \varepsilon B_1(A,\phi) + \varepsilon^2 B_2(A,\phi) + \dots$$

la serie essendo convergente per $|\varepsilon|<\varepsilon_0$, uniformemente su $\bar{V} \times T^{\ell}$. Poiché B è un integrale primo si dovrà avere $\{B,H_\varepsilon\} = 0$, cioè: all'ordine 0 in ε

(1.38) $$\{h(A),\ B_0(A,\phi)\} \equiv \langle\omega(A),\nabla_\phi B_0(A,\phi)\rangle = 0$$

(1.39) $$\{V(A,\phi),\ B_0(A,\phi)\} + \{h(A),\ B_1(A,\phi)\} = 0$$

all'ordine 1 in ε.

Indicando al solito con $B_k^{\nu}(A)$ i coefficienti di Fourier di $B_k(A,\phi)$, la (1.39) implica $\langle\omega(A),\nu\rangle B_\nu^0(A) = 0\ \forall\nu\in Z^{\ell}$, $\nu \neq 0$. Poiché però $\det \nabla_A\omega(A)\neq 0$, se $\nu\neq 0$ abbiamo $\langle\omega(A),\nu\rangle\neq 0$ almeno su un insieme denso in V, e quindi $B_0^{\nu}(A) \equiv 0$ per

$\nu \neq 0$ a causa dell'analiticità. Dunque B_0 non dipende da ϕ : $B_0 = B_0(A)$. Pertanto la (1.39) implica

$$(1.40) \qquad \langle \nabla_A B_0(A), \nabla_\phi V(A,\phi)\rangle - \langle \omega(A), \nabla_\phi B_1(A,\phi)\rangle = 0$$

Pertanto, a meno che non sia $f_\nu(A) = \langle\omega(A),\nu\rangle \hat{f}_\nu(A)$, passando ai coefficienti di Fourier abbiamo che per $\nu \neq 0$: $\langle\omega(A),\nu) = 0 \Leftrightarrow \langle\nabla_A B_0(A),\nu\rangle = 0$.

Dunque per una V generica i vettori $\nabla_A B_0(A)$ e $\omega(A)$ saranno paralleli. $\exists$ pertanto $A \to \lambda\,(A) \in R$ tale che

$$(1.41) \qquad \nabla_A B_0(A) = \lambda(A)\nabla_A h(A) = \lambda(A)\omega(A)$$

Pertanto $B_0(A) = F(h(A))$, $\lambda(A) = F'(h(A))$ per una qualche funzione analitica F. Di conseguenza, per la (1.40)

$$(1.42) \qquad B_1(A,\phi) = V(A,\phi)\, F'(h(A)) + C_1(A)$$

$C_1(A)$ analitica. Possiamo quindi concludere che B ha la forma:

$$(1.43) \qquad B(A,\phi,\varepsilon) = F(h(A)) + \varepsilon F'(h(A))V(A,\phi) + \varepsilon C_1(A) + \varepsilon^2 B_z + \ldots =$$

$$= F(h+\varepsilon V) + \varepsilon C_1(A) + 0(\varepsilon^2) = F(H_\varepsilon) + \varepsilon(B_0' + \varepsilon B_1' + \ldots)$$

Poiché B e $F(H_\varepsilon)$ sono costanti dal moto, così sarà anche la serie $(B_0' + \varepsilon B_1',+\ldots)$, che ha raggio di convergenza non nullo ancora per l'analiticità di B e $F(H_\varepsilon)$. Ripetendo il ragionamento si vede che anche $B_0' + \varepsilon B_1',+\ldots$ è un integrale primo analitico della forma $F_1(H_\varepsilon(A,\phi)+\varepsilon(B_0'' + \varepsilon B_1'' + \ldots)$, e ripetendolo indefinitamente si conclude

$$(1.44) \qquad B = F(H_\varepsilon)+\varepsilon F_1(H_\varepsilon)+\varepsilon^2 F_2(H_\varepsilon)+\ldots\varepsilon^n F_n(H_\varepsilon)+0(\varepsilon^{n+1})$$

Per l'analiticità di B la serie deve convergere:
$B(A,\phi,\varepsilon) = F_\varepsilon(H_\varepsilon(A,\phi))$ per una qualche F_ε il che prova il teorema.

Questo teorema negativo ci dice che il metodo delle perturbazioni canoniche richiede un'analisi più approfondita e sottile se si vogliono trarre delle conclusioni positive. Il risultato positivo è costituito dal celebre teorema di Kolmgorov-Arnold-Moser, il cui enunciato è già di per sé assai sottile.

Teorema (Kolmgorov-Arnold-Moser). Sia $H(A,\phi,\xi) = h_0(A)+\varepsilon f_0(A,\phi), \varepsilon>0$ olomorfa in $W(\rho,\xi, ;V)$ definito come sopra. Sia ancora:

(1.45) $$\sup_{w(\rho,\xi;v)} |\nabla_A h_0(A)| \leq E_0$$

(1.46) $$\|f_0(A,\phi)\|^1_{\varepsilon,\xi} = \sup_{w(\cdot)} (|\nabla_A f_0(A,\phi)|+\rho^{-1}|\nabla_\phi f_0(A,\phi)|) \leq \varepsilon_0$$

(1.47) $$\sup_{W(\cdot)} |(\frac{\partial^2 h_0}{\partial A_i \partial A_j})^{-1}| \leq \eta_0$$

Allora esistono costanti positive $B, C_0, \alpha_1, \alpha_2, \alpha_3$ tali che se

(1.48) $$B(\varepsilon_0 \; C_0 \; (E_0\eta_0\rho_0^{-1})^{\alpha_1} \; (E_0 \; C_0^{\alpha_2} \; \xi^{-\alpha_3} < 1$$

il sistema Hamiltoniano $H(A,\phi,\varepsilon)$ è "canonicamente integrabile" su un sottoinsieme $\Gamma(f_0)$ di $V \times T^\ell$ tale che $\mu(\Gamma(\cdot)) \to \mu(V \times T^\ell)$ per $\varepsilon_0 \to 0$. Più precisamente:

(1) $\Gamma(f_0) = Q(f_0) \times T^\ell$, dove $Q(f_0) \subset V$ è descritto come segue: $\exists G_1$ dipendente solo da ℓ e $K_1(\varepsilon_0) \downarrow 0$ per $\varepsilon \downarrow 0$ tale che:

(1.44) $$Q(f_0) \supset \{A \in V: |\langle\omega(A),\nu\rangle|^{-1} \leq G_1 \; K_1^{-1}(E_0\eta_0\rho_0^{-1})^{-\ell} \; |\nu|^{\ell+\frac{1}{2}}\}$$

e vale la minorazione (notare che A dipende da ε_0 cioè da f_0)

(1.50) $\mu(Q(f_0)) \geq (1-K_1(\varepsilon_0))\mu(V)$, $\mu(\cdot)$ = misura di Lebesgue.

(2) Ogni toro in $\Gamma(f_0)$, cioè ogni punto in $\Gamma(f_0)$ della forma $\{A(f_0)\}\times T^\ell$, $A(f_0)\in Q(f_0)$, è un toro invariante per H_ε, percorso quasi-periodicamente con frequenza $\omega(A(f_0))$.

(3) $\exists\,\ell$ funzioni $\phi_1^1,\dots\phi_\ell^1\in C^\infty(V\times T^\ell)$, cioè $\phi_1\in C^\infty(T^\ell\times V,T^\ell)$ e ℓ integrali primi $A_1^1,\dots,A_\ell^1\in C^\infty(V\times T^\ell)$ tali che (A_1,ϕ_1), (A,ϕ) possono essere fatte corrispondere canonicamente su $\Gamma(f_0)$, la trasformazione canonica essendo generata dalla soluzione dell'equazione di Hamilton-Jacobi: più precisamente $\exists\Phi\in C^\infty(V\times T^\ell;\, R)$ tale che le equazioni

(1.51)
$$A = A_1 + \nabla_\phi\Phi(A_1,\phi)$$
$$\phi_1 = \phi + \nabla_{A_1}\Phi(A_1,\phi)$$

possono essere esplicitate in $\Gamma(f_0)$ per dare $A = A(A_1,\phi_1)$, $\phi = \phi(A_1,\phi_1)$, $(A_1,\phi_1)\in\Gamma(f_0)$, e Φ risolve su $\Gamma(f_0)$ l'equazione di Hamilton-Jacobi:

(1.52)
$$h_0(A_1+\nabla_\phi\Phi(A_1,\phi))+f_0(\nabla_\phi\Phi(A_1,\phi)+A_1,\phi) = h(A_1,\varepsilon)$$
$$(A_1,\phi)\in\Gamma(f_0).$$

Le equazioni parametriche di ogni toro perturbato in $\Gamma(f_0)$ su cui è costante lo integrale primo A_1 sono date dalle (1.51) stesse.

Osservazioni. Non si proverà a dimostrare il teorema KAM , qui enunciato per il caso analitico . Ricorderò soli i tre punti assolutamente chiave.

(1) Partendo da $h_o(A) + f_o(A,\phi)$, si effettua un primo cambiamento di variabile canonico per rimuovere la perturbazione all'ordine ε_o^2, usando la teoria canonica al primo ordine. Si cerca cioè di arrivare all'Hamiltoniana canonicamente equivalente $h_1(A_1) + f_1(A_1,\phi_1)$, che avrà parametri $E_1,\eta_1,\xi_1<\xi_o$ $\rho_1<\rho_o,\varepsilon_1$, con $\varepsilon_1 = O(\varepsilon_o^2)$.

(2) Per ottenere stime ragionevoli sui piccoli denominatori $\langle\omega(A_1,\nu\rangle^{-1}$ è necessario che la perturbazione abbia un numero finito di componenti di Fourier. Pertanto si introduce che il "cut-off di Arnold", cioè si scrive $f_o(A,\phi) = \sum_{|\nu|<N_o} f_\nu(A)e^{i\langle\nu,\phi\rangle} + \sum_{|\nu|\geq N_o} \phi_\nu(A)e^{i\langle\nu,\phi\rangle}, = f_o^I + f_o^{II}$, e si sceglie N_o tale che $\|f_o^{II}\|^1_{\rho_o,\xi_o} = O(\varepsilon_o^2)$. Fatto ciò, si riesce a stimare il volume dei punti "vicini alla risonanza cioè dagli A_1 tali che $|\langle\omega(A_1),\nu\rangle|^{-1}$ $\leq C_1|\gamma|^\ell$ non è verificata.Ciò permette di definire e stimare la funzione generatrice $\Phi_o(A_1,\phi)$ sul complemento di tale insieme, chiamato $\Gamma_1(f_o)$.

(3) Per le stime precedenti si prova che la $\Phi_o(A_1,\phi)$ genera una transf. canonica $C_1:(A,\phi)\longleftrightarrow(A_1,\phi_1)$ su $\Gamma_1(f_o)$ tale che nelle nuove variabili si ottiene $h_1(A_1) + f_1(A_1,\phi_1)$ con parametri come sopra.

(4) Si itera il procedimento (iterazione di Kolmogorov$\Leftrightarrow$metodo di Newton), scegliendo un cut-off di Arnold N_2, e si arriva a $h_2(A_2) + f_2(A_2,\phi_2)$ con $f_2 = O(\varepsilon_0^4)$ in generale:

$h_k(A_k) + \phi_k(A_k,\phi)$, $f_k = O(\varepsilon^{2k})$: procedimento superconvergente.
La condizione (1.49) è quella che assicura il processo iterativo può essere compiuto $\forall k$.

(5) L'insieme $Q(f_0)$ è quindi estremamente complicato, perché ottenuto togliendo ogni volta un aperto di V, e u: esso assomiglia a un cantoriano e in certi casi lo è.

NOTA BIBLIOGRAFICA

Le nozioni di sistema integrabile, canonicamente integrabile, ecc. sono tratte da:

(1) V.I. ARNOLD, Methods Mathématiques de la Mécanique Classique, MIR, 1977.

(2) G. GALLAVOTTI,Meccanica Elementare, Boringhieri, 1986.

Per la dimostrazione del teorema di Liouville-Arnold, Si veda (1).

Per la teoria canonica delle perturbazioni, la forma normale di Birkhoff si veda ancora (1) e (2). Criteri di convergenza della serie di Birkhoff sono stati dati da H. Rüssmann, Math. Ann. 164, 55 (1967) e G. Gallavotti, Comm. Math. Phys. 87, 365 (1982).

Il teorema di Poincaré si trova in H. Poincaré, Methods Nouvelles de la Mécanique Céleste, Vol. I, 1892, Cap. VIII.

Esistono a tutt'oggi molte esposizioni della teoria KAM. Il risultato è stato annunciato da A.N. Kolmogorov, Dokl. Akad. Nank, 98, 527 (1954), e dimostrate da V.I. Arnold nel caso analitico; V. Arnold, Russ. Math. Surveys, 18, 5 (1963) e 18, 85 (1963). J. Moser ne ha ottenuto indipendentemente una dimostrazione nel caso differenziabile: J. Moser, Nach. Akad. Wiss. Göttinger, IIa, 1, 1962.

L'enunciato più completo qui riportato è quello di Chierchia-Gallavotti, Nuovo Cimento, B67, 277, 1982.